Density estimation with mixed model tuning (Old Faithful geyser data)

Old Faithtful; mixed model smooth; bin width 0.02 min.

Old Faithtful; mixed model smooth; bin width 0.05 min.
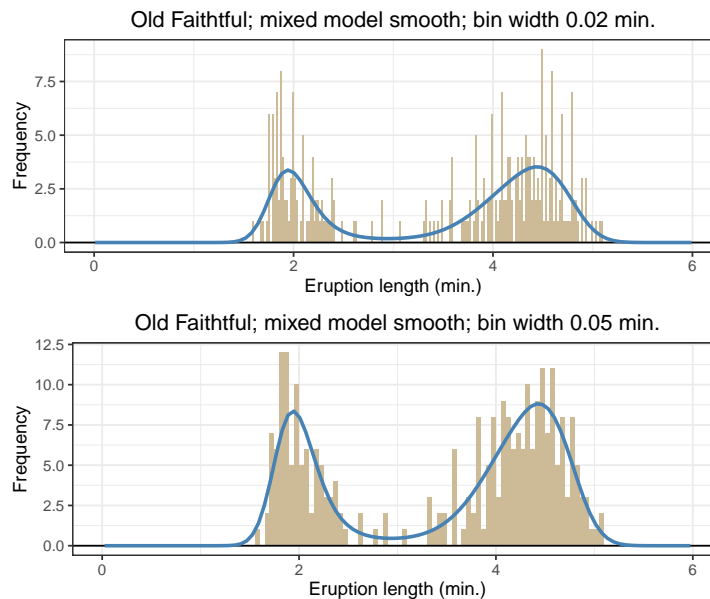
Automatic smoothing of the histogram of eruption times of Old Faithful, using the pure random effects model, for two bin sizes. R code in `f-geyser-mixmod.R`

```
# Plots for optimal smoothing (Old Faithful geyser data)
# A graph in the book 'Practical Smoothing. The Joys of P-splines'
# Paul Eilers and Brian Marx, 2019

library(ggplot2)
library(gridExtra)
library(JOPS)
library(MASS)

# Get the data
data(faithful)
u = faithful[, 1]   # Eruption length
bw1 = 0.05
brks1 = seq(0, 6, by = bw1)
h = hist(u, breaks = brks1, plot = F)
x = h$mids
y = h$counts
Data = data.frame(x = x, y = y)
Dat = data.frame(u = u)
nseg = 20
lambda = 1
d = 3

# Iterative Poisson smoothing, updating tuning based on diff of
# coeffs
aics = NULL
for (it in 1:20) {
    fit = psPoisson(x, y, nseg = nseg, pord = d, lambda = lambda, show = F)
    a = fit$pcoef
    vr = sum((diff(a, diff = d))^2)/fit$effdim
    lambda_new = 1/vr
    dla = abs((lambda_new - lambda)/lambda)
    lambda = lambda_new
    cat(it, log10(lambda), "\n")
    if (dla < 1e-05)
        break
```

```
}

# Gridded data for plotting
Fit1 = data.frame(x = fit$xgrid, y = fit$mugrid)

plt1 = ggplot(Dat, aes(u)) +
  geom_histogram(fill = "wheat3", breaks = brks1)+
  geom_hline(yintercept = 0) +
  xlab("Eruption length (min.)") + ylab("Frequency") +
  ggtitle(paste("Old Faithful; mixed model smooth; bin width", bw1, "min.")) +
  geom_line(data = Fit1, aes(x = x, y = y), col = "steelblue", size = 1) +
  JOPS_theme()

# Second histogram
bw2 = 0.02
brks2 = seq(0, 6, by = bw2)
h = hist(u, breaks = brks2, plot = F)
x = h$mids
y = h$counts
Data = data.frame(x = x, y = y)

nseg = 20
lambda = 1
d = 3

# Iterative Poisson smoothing, HFS tuning of lambda
aics = NULL
for (it in 1:20) {
  fit = psPoisson(x, y, nseg = nseg, pord = d, lambda = lambda, show = F)
  a = fit$pcoef
  vr = sum((diff(a, diff = d)) ^ 2) / fit$effdim
  lambda_new = 1 / vr
  dla = abs((lambda_new - lambda) /lambda)
  lambda = lambda_new
  cat(it, log10(lambda), '\n')
  if  (dla < 1e-5) break
}

# Gridded data for plotting
Fit1 = data.frame(x = fit$xgrid, y = fit$mugrid)

plt2 = ggplot(Dat, aes(u)) +
  geom_histogram(fill = "wheat3", breaks = brks2)+
  geom_hline(yintercept = 0) +
  xlab("Eruption length (min.)") + ylab("Frequency") +
  ggtitle(paste("Old Faithful; mixed model smooth; bin width", bw2, "min.")) +
  geom_line(data = Fit1, aes(x = x, y = y), col = "steelblue", size = 1) +
  JOPS_theme()

# Make and save graph
grid.arrange(plt2, plt1, nrow = 2, ncol = 1)
```